IDC

5 Speen Street Framingham, MA 01701 USA P.508 872 8200 F 508.935.4015 www.idc.com

WHITE PAPER

A New Disruptive IP Storage Technology Paves the Way for Advanced, Cost-Effective, and Protected Storage for Corporate and Consumer Networks

Dave Reinsel
April 2005

EXECUTIVE SUMMARY

Disruptive opportunities can be difficult to identify within any given industry, but they do exist. Defining the right product and the right initial target market that meet the disruptive opportunity's requirements is even more challenging, especially within mature and well-entrenched industries, such as the network storage industry.

Zetera's patent-pending technology has the potential to be a disruptive solution within the network storage industry due to the technology's ability to match or exceed many of the capabilities of today's existing storage solutions with one that is lower-cost, possessing few compromises.

Consistent with companies that have deployed successful disruptive technologies, Zetera has a focused approach. The company is marketing its technology initially within market segments (e.g., SMB) where there is little conflict with established high-end storage systems and within consumer storage solutions where there is an increasing need for inexpensive, easy-to-deploy protected storage. Once the technology is deployed, adopted, and proven successful in these markets, it can be adapted and positioned within established high-end storage array markets.

Nevertheless, companies introducing a disruptive technology such as Zetera's should expect a fair amount of reticence from users who are fully entrenched in today's typical networking infrastructure — there is always a tendency to latch on to proven legacy technologies. First on Zetera's list of priorities should be a strategy to partner with credible vendors (the company has already announced relationships with Bell Microproducts, StorCase, and Netgear) and to achieve success in basic environments to prove the reliability and robustness and, more important, the simplicity of its technology.

Situation Overview: Disrupting the Storage Industry with IP

Disruptive technologies rarely are guaranteed to be successful. Companies introducing potentially disruptive technology must invest much time and effort in defining the right product and the right initial target market that meet the disruptive opportunity's requirements; this is especially challenging within mature and well-entrenched industries, such as the network storage industry. Many times, however, it will be the

Sponsored by Zetera

All these changes mean that the storage capacities and practices IT managers thought would be adequate to support company growth for years are now providing breathing room for just a few months and are causing increasingly frequent disruptions in systems and applications. IT managers want storage solutions that meet their unique business and organizational requirements and scale fluidly to meet tomorrow's anticipated and unanticipated demands

- They want solutions that are easy to set up, allowing simple data migration from existing systems, without requiring any expansion in IT staffing levels
- They want solutions that require low initial investments without sacrificing the ability to expand as needs grow
- They want integrated packages that address specific needs (i.e., server consolidation, improved data protection, email archiving) without disrupting current operations.

Of course, all these "wants" are complemented with fixed IT budgets and head counts and an overarching corporate mandate to improve revenue and profitability

Networked Homes Increasing the Need for Advanced, Cost-Effective Home Storage

Over the past year, consumers have seen an explosion in the range and type of home networking and personal entertainment devices that require expanded levels of storage The home is evolving from a place with one PC accessing the Internet to a broad range of systems (e.g., desktops, digital video recorders [DVRs], game systems, music players) that all have storage.

Today, it is not hard to find households that have almost a terabyte of storage installed across various devices Although these home users are increasingly deploying higher-capacity and larger numbers of drives, they are doing so in an uncoordinated and nonintegrated fashion (similar to many corporate environments where disparate servers are found frequently). As consumers increase their use of applications for digital photos, music, videos, and personal data (e.g., personal finances), they will need storage solutions that make it very easy to accomplish the following goals:

- Add capacity that is self-configuring and accessible by all systems
- Allocate capacity for individual applications (e.g., DVR)
- Enable more robust data protection (RAID-like) across disks that are physically in separate devices

Unlike corporations that require a myriad of features that can be exploited by talented and salaried IT employees, consumers require a minimum set of storage features that can be deployed and managed via self-configuring utilities and intuitive graphical user interfaces. However, like corporations, consumers have an increasing subset of digital information that requires robust protection from hardware failure, corruption, and accidental deletion In fact, consumers concerned with protecting their families' digital heritage will need to protect and to manage data longer than corporations, perhaps even forever.

Data protection cannot be overemphasized Shared HDDs are used more frequently within storage solutions and run the risk of failure; therefore, they require higher-level features such as RAIDs, which protect data from individual drive failure Although the majority of storage environments (desktops or mobile devices) use only one HDD, many enterprise storage systems and an increasing number of home and small business applications require aggregations of HDDs ranging from as few as two to hundreds of drives.

All of this aggregation, however, also adds a layer of complexity and cost because these storage systems require additional functionality to organize and allocate storage capacity (typically referred to as volumes) across multiple HDDs

The main adopters of aggregated storage systems are enterprises that have the need and the ability to deploy and manage these inherently more complex systems The main effort of providing aggregation focuses on the controllers and drives rather than how the user naturally interacts and interfaces with the data In addition, despite the billing of SAN and NAS devices as "network architectures," many designs have various direct-attached legacies. Consequently, improvements in these network storage scenarios have been realized in increments, at great cost, and by brute force

New approaches, such as Zetera's, dispense with the controller and local bus and connect the user and application directly to the network storage fabric. Thus, dramatic improvements are possible without incurring additional cost because the legacy technologies no longer have to be maintained

Strengths and Weaknesses of Existing Storage Systems

Four fundamental factors drive the success of a network storage system. These factors are:

- **Cost.** The total cost to the user to set up, maintain, and use the network storage system
- **Scalability.** The ability to scale the system by adding capacity to the overall system as an enterprise's needs grow
- **Performance.** The system's ability to deliver data to applications and users
- **Reliability.** Ensuring data integrity in the event of a system failure

Any given network storage system may optimize one or some combination of these factors for each target market, but optimizing all factors simultaneously poses a considerable challenge. For instance, low-end NAS systems provide cost-effective storage for home and SMB environments at the expense of performance and, possibly, scalability A SAN provides larger enterprises with high-performance capabilities, but at the expense of cost, because special equipment such as Fibre Channel switches and specialized knowledge is required RAID systems cover various aspects of reliability, where one or even two drives can fail without losing data integrity; however, these systems can suffer in performance due to controller bottlenecks and in cost due to highly complex layers of technology for data redundancy.

THE ZETERA SOLUTION

Zetera Corporation, an Irvine, California–based company, introduced a new technology for network storage based on the industry standard IP. Zetera's technology puts compute power and intelligence on the HDDs themselves so that the storage system's capabilities and functions are handled by the distributed processing on the drives alone. It is designed to meet the core storage requirements of businesses and consumers by leveraging off-the-shelf technology. At the heart of Zetera's protocol is a storage fabric built upon the internetworking protocol standards developed by the Internet Engineering Task Force. This same IP is the foundation of the Internet and just about every private data network.

Zetera's approach isn't "storage over IP" as discussed by many companies in the storage industry. These solutions, as typified by iSCSI, retain the attributes of these legacy storage protocols. For instance, iSCSI builds a tunnel inside IP and moves the entire SCSI protocol. Both end devices are still bound by the functionality inherent in SCSI. In contrast, rather than using IP as an extension cord between clients and storage devices, Zetera uses IP as the storage paradigm itself.

Overview of Zetera's Technology

To understand Zetera's solution, companies must review the basic properties of IP. IP offers a number of interesting properties that developed in response to the scope and complexity of the networks it supports. These properties include:

- **IP addressing.** A system for identifying and self-registering every unique entity on the network.
- **Multicasting.** A system for enabling a single system to communicate simultaneously with multiple systems using IP protocols such as Universal Data Packet (UDP).
- **Broadcasting.** A system for enabling a device to communicate simultaneously with all other devices.
- **Network address translation.** An ability to connect devices on two different networks, providing both greater security and greater scalability.
- **Routing.** A system of intermediary devices that manages the flow of information between communicating devices regardless of failures anywhere in the network.
- **Security.** A set of standards that ensures the authenticity of communications as well as a method (encryption) for preventing unauthorized retrieval of information by third parties.

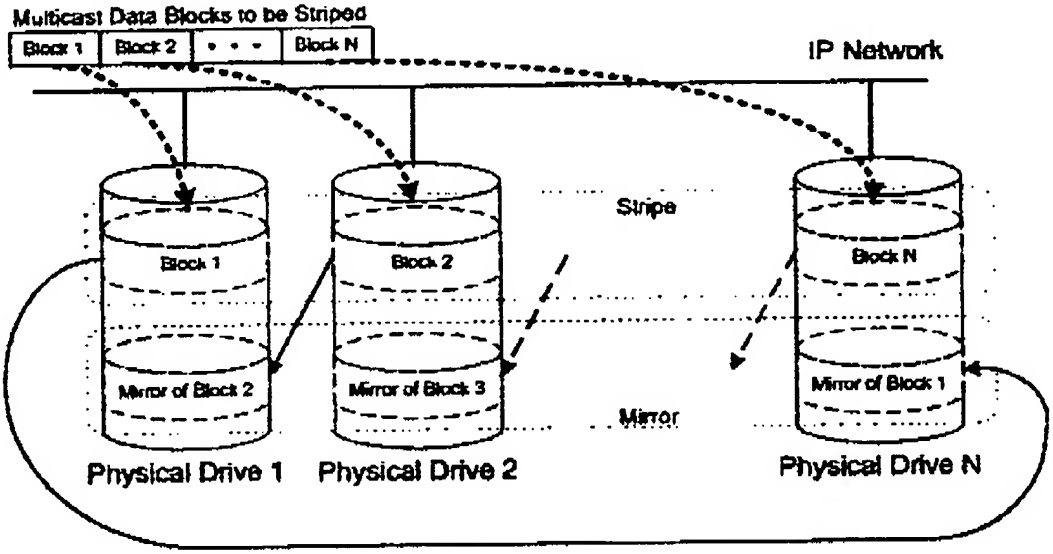
Zetera's solution utilizes these native IP capabilities (e.g., routing, self-configuration, and organization) to provide a virtual data connection that aggregates drive performance as well as enables transparent, seamless capacity expansion, and self-healing on drive or link failures. The company's protocol offers true peer-to-peer storage networking, where any PC can access any storage component, as long as there is an IP connection between them.

Example Application (RAID-10)

As an example, consider how the Zetera technology can be applied to use the fundamental RAID building blocks to create a RAID-10 system. Figure 1 depicts a Zetera RAID-10 system with N drives where each drive has a logical stripe partition and a logical mirror partition, both accessible via IP addresses. The striped partitions combine to form one striped volume, and the mirrored partitions combine to form a mirror of the striped volume.

FIGURE 1

Example of Zetera's RAID-10 System with N Drives



Source: Zetera, 2005

In Figure 1, data is broken into data blocks and sent via multicast across all drives in the logical striped volume as indicated by blocks 1 through N. In addition, the logical stripe partitions then send the data to a mirror partition. The mirrored volume is staggered by one drive so that the mirror partition on drive 1 is mirroring data block 2 from the stripe on drive 2 and so on.

This configuration allows the array to recover from any single drive failure. For instance, if drive 2 fails and drops out of the array, data from drive 2 can be read from the mirrored partition on drive 1 in a manner that is transparent to the user. In addition, when drive 2 is replaced, it can be reconstructed quickly. Drive 2's striped data is recovered by simply copying its mirrored data from drive 1. Also, drive 2's mirrored data is recovered by simply copying the striped data from drive 3.

TCP protocol can also be quite extensive, eating into performance. TCP data ordering is not required because the HDD expects random data ordering due to the scatter-gather nature of the HDD cache. UDP does not have this overhead and allows for line-rate speeds.

Second, performance of a Zetera RAID-based system may exceed that of conventional RAID arrays with similar reliability because parity does not have to be maintained, nor is a RAID controller present to create a bottleneck. Data redundancy is maintained through mirrors and stripes.

Increased Reliability

Reliability for a Zetera-based RAID system may exceed that of a conventional RAID-5. In the preceding RAID-10 example, the RAID system has one stripe and one mirror. Assuming a configuration with three drives having one replication of the data, then one drive can fail without data loss. If there are four drives in the array, then two nonadjacent drives could fail without data loss. Such a four-drive array may be more robust than RAID-5. Reliability through data replication is obviously coupled to scalability; however, customers can decide for themselves if the incremental cost is worth the reliability.

Easy Manageability

Management of Zetera-based network storage solutions can be performed by any individual with basic networking experiences and the ability to place a drive in an enclosure. Enterprises already have onsite network administrators who can handle deployment and maintenance of a Zetera solution. Use of IP addresses and multicasting allows a Zetera system to be easily placed anywhere on a network that offers maximum flexibility in a complete networking solution.

Market Opportunities

The Zetera technology addresses the four critical storage solution criteria — cost, scalability, performance, and reliability — as effectively as, if not more effectively than, many existing network storage solutions.

Large businesses have IT departments that have the ability to deploy and to manage complex storage environments or at least have the resources to purchase the expertise, many small and medium-sized businesses do not. Yet, many small businesses have storage needs beyond what single disk drives can provide but cannot afford the high-priced storage solutions from well-established storage OEMs. Moreover, the skills necessary to manage such solutions are typically not found within smaller SMBs. Zetera's technology offers an acceptable alternative to costly storage solutions with a low-cost, easy-to-deploy, and easy-to-manage integrated storage system.

As noted earlier, IDC also sees consumers becoming important purchasers of storage over the next few years as the need to protect data increases with the growing amount of "precious data" (e.g., personal video, family photos) being stored onto disk drives. In addition, consumers continue to add discrete islands of storage to the home environment via multiple devices such as MP3 players, PVRs, multiple PCs, and external storage.

Zetera appears to be on a solid path to address all the criteria of a disruptive technology. The company provides the basis for replicating essential building blocks for network storage such as RAID-0 stripes and RAID-1 mirrors. In addition, the technology may exceed the reliability of existing systems through an advantageous incremental cost route for consumers by placing fine-grained control in the hands of the customer. Finally, by pursuing the SMB market where legacy systems are not entrenched, Zetera may build a foundation from which it could have the opportunity to migrate up through more established markets.

Technology Licensing

The cost benefits of Zetera's solution are noteworthy (e.g., elimination of RAID controllers, TOEs; HBAs, custom switches; single, simplified HDD interface); however, it is significant that Zetera's controller technology, while leveraging industry-standard technology, is itself proprietary and will require licensing by industry participants.

Few companies have been successful with large-scale licensing and adoption of proprietary technology. There are many examples: Beta versus VHS; 1394 versus USB, Memory Stick versus CF, MMC, or SD; and the list goes on. That is not to say that it can't be done. For example, Dolby, Minicassette, CDMA, and Rambus are proprietary technologies that have been deployed successfully. The fact that Zetera is leveraging one of the most successful and widely adopted technologies, notably IP, will make its task less daunting than most.

Other challenges may include the following:

- The need to work with Microsoft to ensure that there are no interoperability issues. The protocol is nonstandard and hence not part of the Microsoft-certified world. If Windows is a target market, then Microsoft needs to be accommodated.
- To penetrate established storage environments, some sort of data migration path needs to be established to aid the movement or migration of data.
- HDD vendors may view the Zetera technology as "just another interface to support." Such a prevailing attitude would need to be addressed quickly and thoroughly.

Conclusions and Final Thoughts

The changing information needs of businesses and the advent of the networked home with its growing need for storage capacity and data protection will dramatically change the ways companies and consumers buy and use storage in the coming years.

Zetera, with its fully IP-based storage solution, is in a strong position to capitalize on these changes and become a disruptive technology in the network storage market. The company's use of IP addressing and multicasting to access virtual drives spanning multiple physical drivers matches many existing network solution capabilities and has the potential to exceed many solutions in terms of cost, scalability, performance, and reliability.